

18th ICCRTS

Title of Paper: Time Series Modeling of Army Mission Command Communication Networks: An Event-Driven Analysis

Topics: Topic 8: Networks and Networking; Topic 5: Experimentation, Metrics, and Analysis; Topic 6: Modeling and Simulation

Authors:

Laura R. Marusich
U.S. Army Research Laboratory
Aberdeen Proving Ground, MD 21005

Norbou Buchler
U.S. Army Research Laboratory
Aberdeen Proving Ground, MD 21005

Point of Contact:

Laura R. Marusich
Human Research and Engineering Directorate
U.S. Army Research Laboratory
Aberdeen Proving Ground, MD 21005
410-278-8635
laura.r.marusich.ctr@us.army.mil

Abstract: We examine the communication time series of a fully-networked Army coalition command and control organization. The network comprised two echelons of command at the Division and Brigade levels over a two-week military scenario exercise. We used time series analysis to predict the communications record based on an external work variable of the number of important scenario events occurring across time. After taking into account structural features of the time series we examined the remaining variability in email and phone communication. We found that the exercise scenario events were not a significant predictor of the Divisional communications, which were best fit by an auto-regressive model of order 1. The occurrence of scenario events, however, did predict the Brigade communication time series, which were well fit by a lag dependent variable model. These results demonstrate that Brigade communications responded to and could be predicted by battlefield events, whereas the Division communications were only predicted by their own past values. These results highlight the importance of modeling environmental work events to predict organizational communication time series and suggest that network communications are perhaps increasingly dependent upon battlefield events for lower echelons of command closer to the tactical edge.

Report Documentation Page			Form Approved OMB No. 0704-0188		
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE JUN 2013		2. REPORT TYPE		3. DATES COVERED 00-00-2013 to 00-00-2013	
4. TITLE AND SUBTITLE Time Series Modeling of Army Mission Command Communication Networks: An Event-Driven Analysis			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S)			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Laboratory, Human Research and Engineering Directorate, Aberdeen Proving Ground, MD, 21005			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES Presented at the 18th International Command & Control Research & Technology Symposium (ICCRTS) held 19-21 June, 2013 in Alexandria, VA.					
14. ABSTRACT We examine the communication time series of a fully-networked Army coalition command and control organization. The network comprised two echelons of command at the Division and Brigade levels over a two-week military scenario exercise. We used time series analysis to predict the communications record based on an external work variable of the number of important scenario events occurring across time. After taking into account structural features of the time series we examined the remaining variability in email and phone communication. We found that the exercise scenario events were not a significant predictor of the Divisional communications, which were best fit by an auto-regressive model of order 1. The occurrence of scenario events, however, did predict the Brigade communication time series, which were well fit by a lag dependent variable model. These results demonstrate that Brigade communications responded to and could be predicted by battlefield events, whereas the Division communications were only predicted by their own past values. These results highlight the importance of modeling environmental work events to predict organizational communication time series and suggest that network communications are perhaps increasingly dependent upon battlefield events for lower echelons of command closer to the tactical edge.					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Same as Report (SAR)	18. NUMBER OF PAGES 47	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

Time Series Modeling of Army Mission Command Communication Networks: An Event-Driven Analysis

1. Introduction

Advances in information and network technology continue to transform the way organizations communicate and operate, so much so that the emergence of networked forms of organization lies at the core of the economic, military, political, and social fabric of the 21st century (Castells, 1996). In a networked organization, the number of potential collaborations is virtually limitless, as is the availability of information. While this can present a number of challenges, operating in such a broadly collaborative and information-rich environment has the potential to confer unprecedented advantages in directing and responding to work events (National Research Council, 2005).

Networked organizations are increasingly dependent upon robust human interactions to enable timely communication and decision-making, particularly in response to critical collaborative work events. This is perhaps most evident in military command and control networks operating in adversarial, time-stressed battlefield work environments, but also applies to critical event responses within health care networks, emergency response networks, corporate networks, and energy infrastructure networks, among others. As in Monge and Contractor (2001), we refer to this type of network as a *communication network*, which defines the behavior of networked organizations by the flow of messages among communicators across space and time. We take a time series analysis approach to understanding patterns of communications in an organization using data collected from a study of a real-world Army command and control (C2) mission environment. Our goal is to determine how well communication may be predicted by the number of important environmental events occurring over time (i.e., an external work variable). Environmental events drive the work of the organization, which manifest as communications. Understanding the diffusion of information is central to work-directed organizational effectiveness (Wang et al. 2011) and constitutes a major research area that addresses timely real-world challenges with important potential applications.

1.1. Communications Time Series – Approaches

Giving weight to discrete environmental events as predictors is a unique approach to the study of communication time series. Barabási and colleagues have examined the time series of inter-message intervals at the individual level, showing that human communications are punctuated by periodic bursts of rapidly occurring events followed by long periods of inactivity, which they argue arises from an internal priority queuing process (Barabási 2010; Vazquez et al. 2006). The broad social science literature, however, is focused on network structures and how they facilitate and constrain human dynamical behaviors (for a review, see Kilduff and Brass 2010). Structural approaches use graph theoretic approaches to map patterns among complex human network interactions. Social network analysis may be used to examine relational and structural patterns in networks and is a widely-applied and increasingly influential technique. Monge and Contractor (2001) identify as many as nine theories and a multitude of mechanisms that have been used to explain the dynamics of organizational networks using social network analytic approaches.

Understanding how an organization functions, however, also requires examining the environmental context of the organization (see Ancona and Caldwell 1992). This argument is not new. In a seminal book, *The Sciences of the Artificial*, Nobel laureate Herb Simon (1969) offers the parable of an “ant wandering on a pebbled beach” to suggest that the apparent complexity of human behavior might be due to environmental factors:

“An ant, viewed as a behaving system, is quite simple. The apparent complexity of its behavior over time is largely a reflection of the complexity of the environment in which it finds itself. I should like to explore this hypothesis with the word ‘man’ substituted for ‘ant’.”

Theorists in the ‘grounded cognition’ movement (Clark 1997; see Barsalou 2010 for a review) have recently adopted this argument by noting that much of the dynamics of human behavior stem from reciprocal causation links to the environment as well as to other humans (e.g., economic systems), and thus the environmental context must be fully considered. We take this perspective here. We use time series analysis techniques to examine the record of email and phone communications in an Army command and control network to discover the extent to which we can predict the organizational communications time series given a second time series of significant work events.

1.2. Time Series Data and Regression Analysis

The general purpose of regression analysis, whether used in time series or cross-sectional data, is to determine the existence and form of the relationships between variables. Regression analysis typically shows how some dependent variable is affected by changes in one or more independent variables.

In the field of marketing, this explanation of one variable in terms of others is known as an *empirical response model* (Parson and Schultz 1976). The overwhelming majority of these models use sales or market share as a dependent variable. Companies construct sales response models to determine what factors (e.g. advertising) influence or drive their sales and to plan marketing strategies accordingly. Sales is the most direct outcome measure of marketing actions, so models with sales as the dependent variable are very common (see Assmus et al. 1981 for a meta-analysis).

In our examination of networked communications, we use the timing of work events as an independent variable (analogous to advertising in the marketing example), and the communications time series as the dependent variable (analogous to sales). A key consideration is that the network communications reflect both the work process itself and the dissemination of both intermediate and final work products as they occur across the communication network over time.

There is some evidence to suggest that organizations increase communications in response to critical events. In a detailed analysis of the email corpus of the Enron Corporation, Diesner and Carley (2005; see also Murshed et al. 2007) found that during the crisis period of financial insolvency, the volume of communications intensified among employees, becoming more diverse with respect to established contacts and formal roles. The Enron crisis is instructive as a

network with a critical period of failure. Other researchers have also found dramatic changes in email usage following major (typically singular) environmental events in organizational networks, such as after a corporate merger (Danowski and Edison-Swift 1985) or after downsizing (Shah 2000). Remaining unexamined is whether communications can be understood in response to more routine and recurrent work events, as opposed to rare crisis events. In our study, the work events are many and varied, and they are formally established as part of a scenario in a military training event exercise. Social network analysis and time series analysis techniques were used to determine the characteristics of communications in the hierarchical network (i.e., a Division and subordinate Brigade staff) during the scenario-based training event. We seek to determine whether the communication time series can be explained and/or predicted on the basis of a time-line of work events.

2. Method

The communications and scenario-event data described in this paper were collected at a two-week U.S. Army simulation-based training event. The Mission Command Battle Laboratory at Fort Leavenworth, Kansas conducted a joint experiment/exercise— with aspects of both an experiment and a training exercise— focused on the operations of the mission command staff composed of a U.S. Division headquarters (n=51) and subordinate U.K. Brigade headquarters (n=28). The network architecture and digitized nature of the event allowed examination of staff communications in a distributed, network-enabled coalition environment. The participants were active duty Soldiers and officers, organic to their military unit. Below we describe the defining characteristics of this military ad-hoc organization, and of the tasks they were required to complete.

2.1. *Defining Characteristics of the Organization*

- a. *Real organization*: Several groups participated in the exercise, including a representative command and control headquarters of a U.S. Division, a fully-staffed U.K. Brigade, and two partially-staffed U.S. Brigade Combat Teams. The participating organizations were existing units, whose staff execute differentiated, well-specified, but also interdependent roles. The units operated in a distributed-fashion over a communications network using specialized military command and control hardware and software.
- b. *Convened to accomplish a particular training mission*: The military organization was staffed and convened specifically to execute and accomplish a particular two-week long training mission.
- c. *Common mission*: Members are interdependent in their work and decision-making about how to proceed with the mission. The organization functions as a *purposive social system*, where members are readily identifiable to each other by role and work interdependently to accomplish one or more collective objectives (Hackman 2008; Hackman and Katz 2010). The responsibility for performing the various tasks and sub-tasks necessary for mission success is divided and assigned among the staff.

2.2. Defining Characteristics of the Tasks

The training scenario in a military exercise generates many overlapping series of event-driven tasks, the resolution of which requires a high degree of coordination among the participating command and control staff. The training scenario involved a coalition environment with a U.K. Brigade operating under the command of a U.S. Division, focused on coalition interoperability. Two U.S. Brigades also operated under the command of the U.S. Division, but neither was fully staffed nor fully engaged by scenario events. The scenario focus was on the networked organizational communications within and between the U.S. Division and U.K. Brigade. Researchers have long pointed out that the nature of a task has a great influence on the steps and processes a group uses to perform the work (e.g. McGrath and Kravitz 1982; Roby and Lanzetta 1958). The tasks of groups in the military domain considered here have four distinguishing features:

- a. *Specific Presenting Problems:* The military command and control staff is tasked with addressing specific problems that occur in the unit's area of operations (AO). The military staff organization must monitor key events and successfully plan and coordinate an effective response, given limited resources. The presenting problems may be kinetic events, such as an improvised explosive device, or civil-military in nature, such as responding to a civil demonstration. At other times, the presenting problem may be a time-sensitive intelligence report of enemy activity that needs to be analyzed and corroborated. At any given time, the organization must coordinate a response to many such presenting problems.
- b. *Adherence to specific tactics, techniques, and procedures:* The groups adhere to formalized military work routines and processes that are known in advance, delegating specific work responsibilities for the various sub-groups and individual members.
- c. *Addressed immediately:* The group operates in an urgent, time-sensitive work environment and is required to immediately coordinate responses to work events that may have adverse cascading effects if not addressed in a timely manner.
- d. *Results in collaborative work products that need to be coordinated and disseminated:* The group is expected to construct specific, detailed material products that will exist independently of the group process or the individual members themselves. For instance, the Commander and his command elements require regular reports from the staff in order to achieve situational awareness of the battlefield environment. The work process itself and the dissemination of both intermediate and final work products occur across the communication network over time.

2.3. Data Collected

Communications: Telephone (VoIP) and email were the two primary methods of communication between staff members during the exercise. For each phone call made and email message sent in our dataset, three pieces of information were automatically electronically logged: the sender, the receiver, and the time of the communication's initiation.

Scenario Events: We also collected information about the various scenario events, including: the time at which each scenario event occurred and the group that was affected by each event. The specific timing of the significant events was known to the researchers ahead of time, but not to the military units participating in the training exercise. Knowing the scenario event timeline is advantageous to researchers interested in examining the correspondence between specific work events (i.e. presenting problems) and the time series of communications.

2.4. Data Analysis – Social Network Analysis

The first step in analyzing this communications dataset was to construct and visualize it as a directed network, shown in Figure 1. Each node in the network represents an individual, and each link represents at least one communication event between two individuals during the exercise. The individuals belonging to the U.S. Division are blue, and those belonging to the U.K. brigade are red. There were 51 individuals in division roles and 28 in U.K. brigade roles, comprising the total 79 nodes in the network.

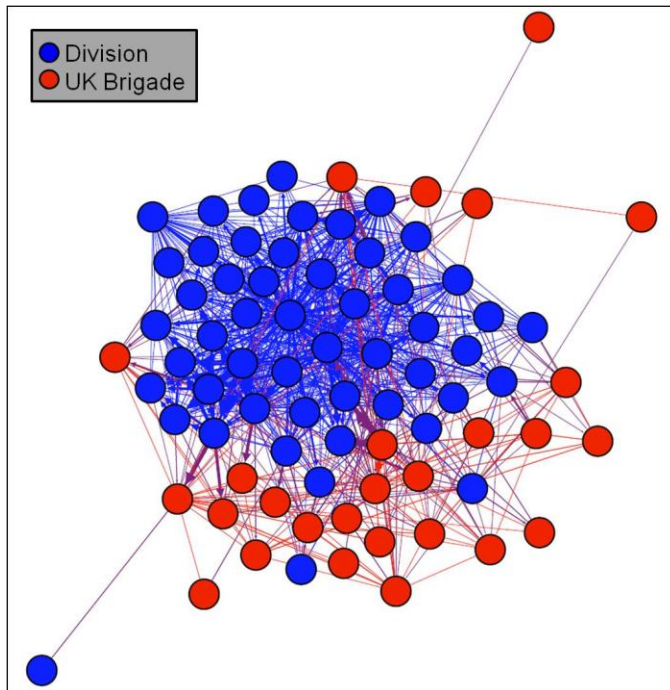


Figure 1. Directed network of communications within and between the U.S. Division and U.K. Brigade.

While each individual in this network was free to communicate with any other, an important feature of the network is that it is composed of two groups differing in both echelon (the brigade is a subordinate element of the division) and nationality (U.S. vs. U.K.). Thus we are interested in both the overall number of links and whether those links connect individuals in the same or different groups. In Figure 2, the network is broken out to depict the internal and external links. The figure shows more internal than external communications and denser internal communications in the Division group than in the U.K. brigade.

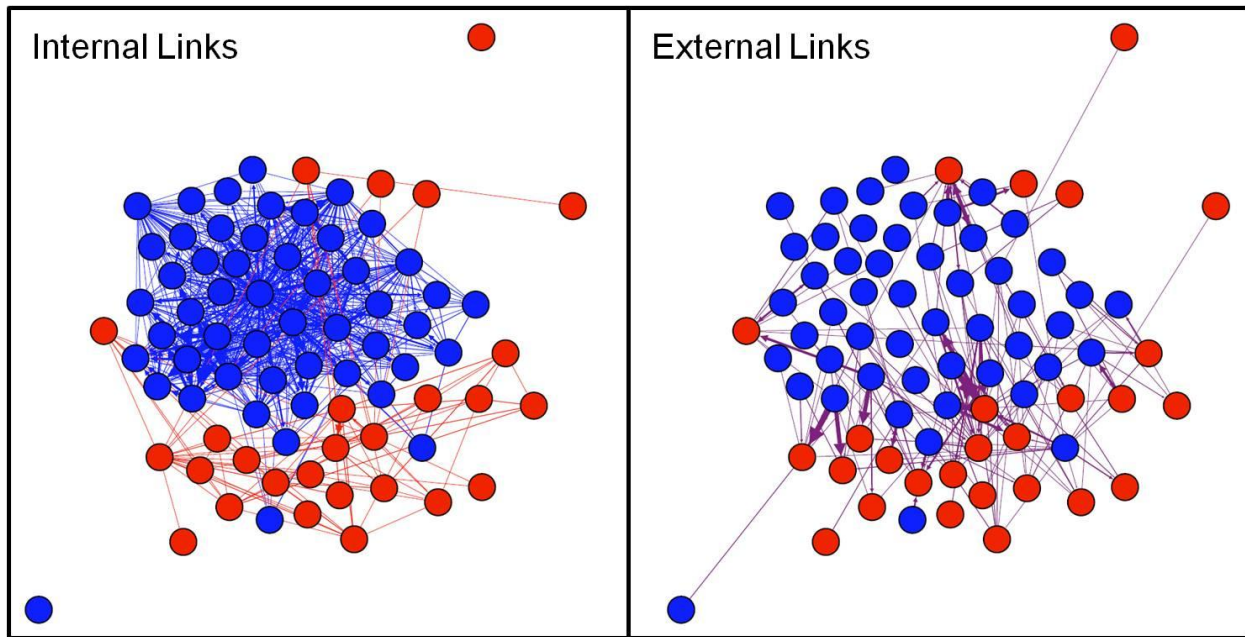


Figure 2. Internal links within the Division and Brigade, and external links between the two groups.

2.5. Data Analysis – Time Series Analysis

The above analyses provide an interesting depiction of the overall distribution of communications within and between groups in the U.S. Division and U.K. Brigade; however, they do not address time. Analyzing the communication events and network statistics as time series allows us to explain the changing dynamics of the collaboration and possibly forecast future communication patterns.

The time series analysis first required aggregating the data into discrete temporal intervals. The appropriate interval size over which to integrate depends on the overall time span and resolution of the data in question. We chose one hour intervals to allow for both a sufficient amount of variability between intervals and a large enough number of time points ($t = 48$) while avoiding too many empty intervals. Thus, we aggregated three variables into intervals of one hour for each of the two groups:

- 1) External Links In - the number of links directed from an individual outside of the group to an individual within the group.
- 2) External Links Out - the number of links directed from an individual within the group to an individual outside of the group.
- 3) Internal Links – the number of links directed from an individual within the group to another individual within the group.

These three variables are plotted against time for the two groups in Figure 3. Also noted on this figure is the time point that marks the start of the second week of the exercise, as well as the five time points during which there was a planned crash in the computer network. One effect of this crash was to prevent all electronic communication; however, phone communications were still

possible. Below these three communication variables, we plot the time series (aggregated by hour intervals) of the number of scenario events that affected both groups.

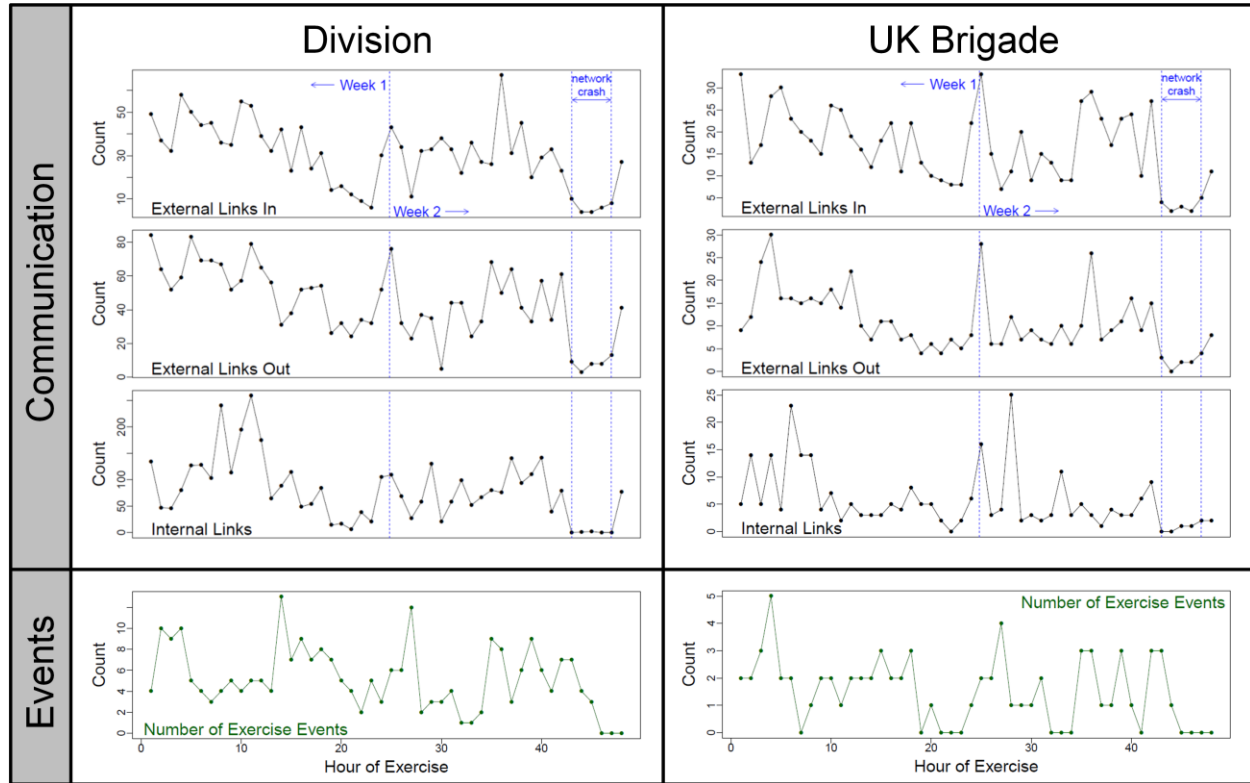


Figure 3. Time series of communications and scenario events for the U.S. Division and the U.K. Brigade over 48 hours of the military exercise.

We are interested in predicting the fluctuations in volume of communication links using the number and timing of the exercise events. However, time series data require additional considerations than are typically necessary for a straightforward regression analysis with cross-sectional data. Random sampling from the larger population is assumed in typical cross-sectional data, such that each case is independent of the others. In time series data, sequential observations of a single case are collected. Instead of each observation representing a random draw from a larger population, the sequence of observations represents one possible realization of a stochastic data generating process. The corresponding assumption to random sampling when using time series data is stationarity: the expected value of the variable of interest must be constant over time.

Assuming the variables are stationary, the sequential observations that are collected in time series data are typically not independent, but are often serially correlated. One of the assumptions required by Ordinary Least Squares (OLS) regression is that the residuals be independent of each other (i.e., the magnitude of one error must not be predictable from the others). If there is serial correlation in the dependent variable, and it is not accounted for by the explanatory variable(s), this assumption will be violated. As a result, past values of the dependent variable are often included as predictors in time series analysis models.

2.6. Stationarity: Trending and Structural Breaks

As described above, one of the key conditions of time series analysis is that the variables be stationary. Figure 3 suggests that our time series data may meet this assumption. Many of the sequences appear to display a downward trend with time. In addition, two features of the exercise seem likely to exert an influence on the values of the dependent variables. The first is the transition between Week 1 and Week 2 of the exercise. The second is a nearly five-hour time period during which a planned computer network crash occurred. Both of these events might be expected to shift the mean of the dependent variable (DV) to a different value at the times they occurred, which is a violation of stationarity known as a *structural break*. Trending and structural breaks can be included in a model along with exogenous variables; however, they can artificially improve fit statistics, and so it is often more appropriate to remove their effects before fitting a model. This is done by simply regressing the DV on these variables, and using the residuals from this regression as the new DV of interest. This new DV will have a mean of zero, and if there were no other factors causing nonstationarity, will be stationary.

Accordingly, we regressed our three communication variables on a linear time variable and two dummy variables, one representing the transition between weeks and one representing the network crash. The residuals from this regression (Figure 5) represent the variability in the communications variables that are not explained by a linear trend with time or the two structural variables.

3. Results

3.1. Time Series Models

There are many varieties of models to fit time series data. We were guided first by theory in choosing among these options (i.e., what is expected to affect the volume of communications in this exercise) and second by parsimony. Our goal in fitting time series models to the data was to determine whether the volume of communications could be predicted based upon previous communications and/or the dynamic volume of scenario events encountered, and whether communications are predicted differently in the U.S. Division than in the U.K. Brigade. We started by examining models that included up to two lags of the DV and the IV, as it is plausible that events and communications from two hours previous could be predictive of communications in the current hour. Terms that were not significant for any of the DVs were removed, and the Akaike Information Criterion (AIC) was assessed as a measure of model fit. In the end, the simplest model that provided the best fit for most of the DVs in the UK Brigade was a Lag Dependent Variable model with one lag of the DV (communication) and only the contemporary values of the IV(scenario events) included as predictors. In the Division, the best model was an autoregressive model of order 1. These two types of models are described in detail below.

Lag Dependent Variable Model

The Lag Dependent Variable (LDV) model differs from simple linear regression in that it includes one or more lags of the dependent variable in its set of predictors, meaning that previous

values of the dependent variable are used to predict the current value. With one independent variable and a single lag of the dependent variable, it takes the following form:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta x_t + \varepsilon_t$$

The inclusion of the lagged terms also makes interpretation of the coefficients on the independent variables somewhat more complicated than simple linear regression where the interpretation of the slope parameter is straightforward; it is the immediate change in the value of y predicted by a one-unit change in x .

An analogous understanding of the relationship between variables in the LDV can be obtained by calculating the short and long run effects of the IV upon the DV. The short run effect is the immediate change in the DV caused by a temporary one-unit increase in the IV. If x were to increase by 1 for a single time point and then return to its prior value in the next, the short run effect is the corresponding immediate change in y . It is equal to β , the coefficient on the independent variable. The long run effect is the asymptotic or eventual change in the DV caused by a permanent one-unit increase in the IV. If x were to increase by 1 and stay at that new value for all future time points, the long run effect is the corresponding eventual change in y . It is calculated by dividing β by one minus α_1 , the coefficient on the lagged dependent variable. Analogous to the intercept term in a simple regression, the equilibrium of the process is calculated by dividing α_0 by one minus α_1 .

Figure 4 depicts the relationship between the two coefficients and the short and long run effects in the LDV. The top panel is a plot of the independent variable over time, with one transient one-unit increase and one permanent one-unit increase. In the middle panel, the corresponding changes in the dependent variable over time are plotted, using a constant value for β and three different values of α_1 . In the bottom panel, α_1 is held constant, and β is allowed to vary. This figure shows that β determines the amplitude of both the short and long run effects, while α_1 determines the time course of the effects, as well as the amplitude of the long run effect.

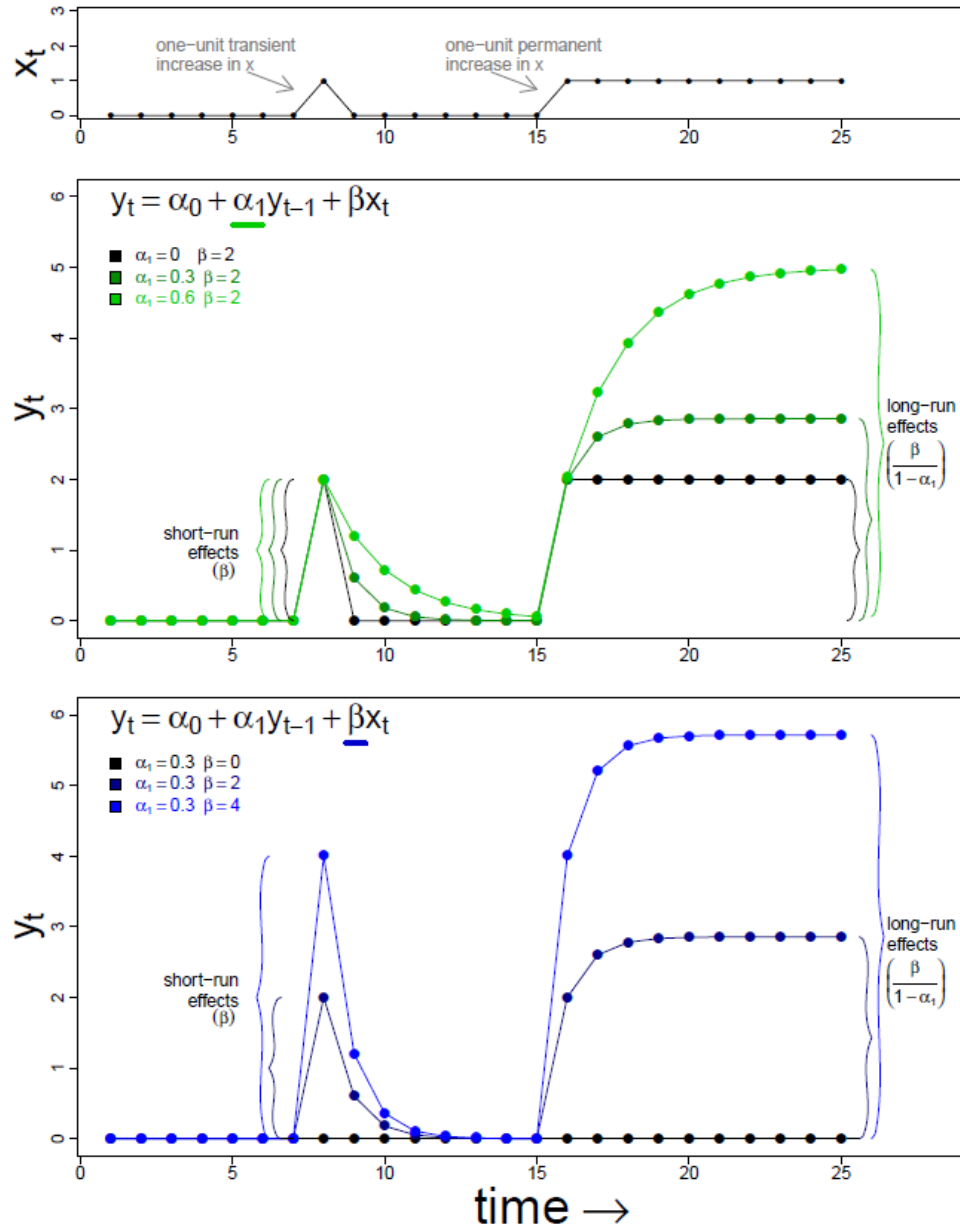


Figure 4. Short- and long-run effects in the Lag Dependent Variable model.

Autoregressive Moving Average Model:

We did not find significant contributions of the scenario events to the communication variables, whether contemporary or lagged values were used in the U.S. Division data. There were, however, significant autocorrelations in the dependent variables. As a result, we found that fitting an autoregressive moving average (ARMA) model to the data provided a good explanation of the dynamics. An ARMA(1,1) model takes the general form:

$$y_t = \alpha_0 + \sum \alpha_i y_{t-i} + \sum \phi_i \varepsilon_{t-i} + \varepsilon_t$$

No external regressors are included, and so this model can be understood as only predicting future values based upon internal dynamics.

We used the Box-Jenkins method for selecting the appropriate ARMA model, finding that the Division data is best explained by an AR(1) model, meaning that it includes one lag of the dependent variable and no moving average term. This takes the form:

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \varepsilon_t$$

In this model, the dependent variable is predicted only from its own previous behavior, and because the model is AR of order 1, only the immediately previous value of y is used to predict the current value. In contrast to the LDV model, the understanding of short and long run effects of the independent variable upon the dependent variable does not apply in this model because there are no external independent variables included as predictors. The process equilibrium is calculated by dividing α_0 by one minus α_1 .

3.2. Model Predictions

The best model of the UK Brigade communications is the LDV, while the best model of the Division communications is an AR(1) model. The de-trended communication time series and the predictions of the best model for each group are plotted in Figure 5. The specific parameters chosen for the two types of models depend on the specific communication variable being studied and are shown in Table 1.

Communication Variable	Division – AR(1)		UK Brigade – LDV		
	α_0	α_1	α_0	α_1	β
External Links In	0.11	0.31*	- 0.30	0.25*	2.53***
External Links Out	0.34	0.34*	0.12	0.18	2.00**
Internal Links	0.60	0.43***	0.08	- 0.01	0.14

Table 1. Model parameters (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$)

The intercept (α_0) values in both models are not particularly meaningful by themselves, considering they reflect de-trended, rather than raw data. In the AR(1) model, the α_1 values reflect the amount of influence the value of the previous time step has on the current one. These values are significant and in the range of approximately 0.3 to 0.4 for all three of the communication variables. The predicted volume of communication at any given point in time resembles a memory growth/decay function that goes up or down depending on the immediately prior communications volume with a step-size of roughly 30% to 40% of that prior value.

The α_1 and β parameters of the LDV model are easier to interpret in terms of long and short run effects. The computed values for these effects are reported in Table 2.

Communication Variable		Short Run Effects	Long Run Effects
External Links In	2.53 ***	10.71 **	
External Links Out	2.00 **	8.85 **	
Internal Links	0.14	0.05	

Table 2. Short and long run effects in LDV model of U.K. Brigade communications

For external links in and external links out, the short run and long run effects are significant. The interpretation of these effects is the following: a temporary increase of one additional scenario event is expected to lead to 2.53 additional links directed into the U.K. Brigade at that time point, and 2.00 additional links directed out of the Brigade. A permanent increase of one additional scenario event is expected to lead to 10.71 additional links in and 8.85 additional links out.

In the AR(1) model of the Division communications data, the internal links show the largest AR coefficient, meaning that the previous value of the volume of links exert a larger influence on the current value. Conversely, in the LDV model of the U.K. Brigade data, the internal links variable is the only one without a significant short or long run effect of scenario-events. The previous value of the number of internal links also did not significantly predict the current value in the U.K. Brigade. This is clearly seen in Figure 5, where the model predictions are capturing only the mean and not the fluctuations over time.

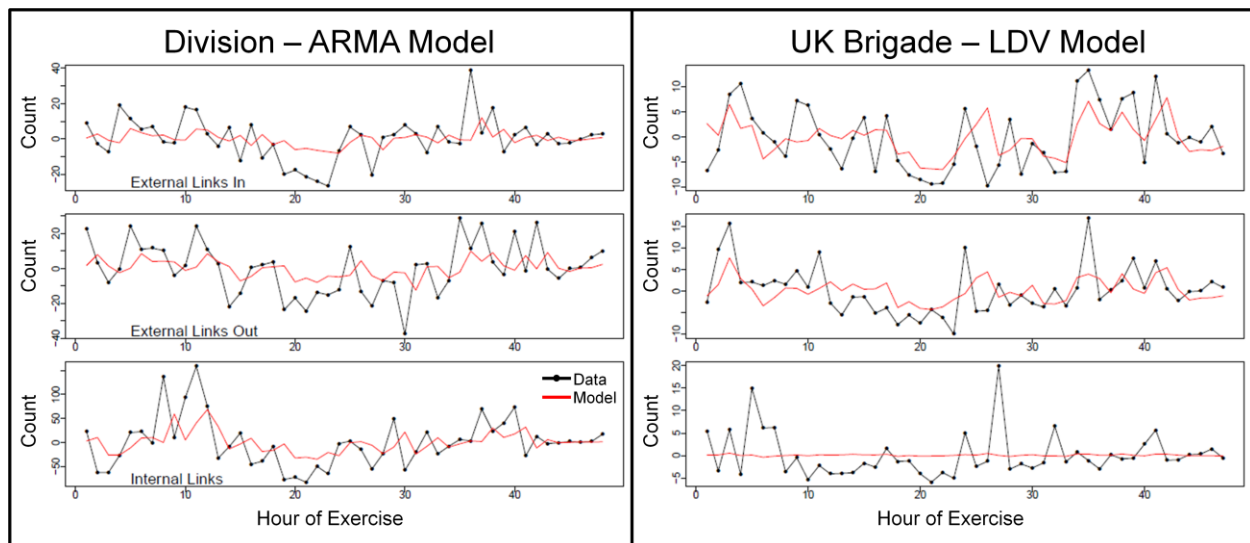


Figure 5. Detrended data (filled black circles) and model predictions (red lines)

4. Discussion

We found that the exercise scenario events were not a significant predictor of the Division communications, which were best fit by an auto-regressive model of order 1, meaning that the best predictor of the volume of communications at a given time point was the volume of communications on the immediately preceding time point. The occurrence of scenario events, however, did predict the external Brigade communication time series (links-in and links-out),

which were fit well by a lag dependent variable model. These results demonstrate that Brigade communications responded to and could be predicted by battlefield events, whereas the Division communications were only predicted by their own past values.

The ability to model and predict the volume of communications of the two groups with any degree of success is in itself interesting, apart from the actual form and interpretation of the models. The data were generated from a complex network of many individuals with distinct roles and the potential to be affected by many different varying factors, both recorded (e.g., scenario events) and unrecorded. There is no guarantee that the communications on one time point would be a significant predictor of the next time point, but in both groups (Division and Brigade), the best models include the number of communication links on the previous time point as a predictor. The number of communication links at each time point was not independent of the others, but were serially correlated (i.e., high numbers of communication links tended to be followed by high numbers of communication links, and vice versa).

In addition, the Brigade data were well predicted by the number of scenario events occurring within the same time period. This is also interesting because the scenario events varied greatly in terms of severity and the nature of the response required, and it is not obvious that a simple measure like the number of events could predict the communications. Using environmental events to predict network dynamics is not a standard approach in the social network analysis literature. Typical goals for time series analyses of network behavior are to identify and visualize significant changes in a network. Models of change often focus on the node level, where changing states of individual nodes leads to changes in the structure of the network over time (e.g. Carley 2003; Snijders 2005). In contrast, we focus on understanding communication networks in the aggregate, and our results suggest that this approach could be a valuable tool for understanding network dynamics.

Prior to the analysis, the time series data were de-trended using a linear time variable and by removing the effects of two structural breaks (the transition between scenario weeks and the network crash period). This preprocessing step was performed to meet assumptions of stationarity – a variable that trends is not stationary, as its mean is changing with time. Similarly, stationarity is not met when the mean of a variable shifts to a new level at a particular time (e.g., the communication data during the network crash). From this perspective, the three variables that we used to de-trend the data could be considered merely nuisance variables.

However, the significant effects of these three variables upon the communications time series provide some interesting information about our data. For example, the communication variables showed significant decreases with time. This suggests the possibility of a type of training effect, where the participants initially established communication links with many different individuals, but pruned these connections over time as they learned which links were most efficient for accomplishing their goals. As a result, the number of internal and external links decreased over time. Secondly, there was a significant increase in communications at the transition between weeks, indicating a type of reset, where the steady decrease in links over time was temporarily offset after a break and change in the mission of the exercise. Finally, there is a significant decrease in communications during the period when the computer network crashed. This is an

expected effect and provides a good check that our time series analysis methods are sensitive to the dynamics of this exercise.

5. Conclusion

One of the primary findings from our analyses is that scenario events can be used as a predictor for the external communication time series from the U.K. Brigade but not for the communication time series from Division headquarters. This finding suggests that as the focus of analysis moves further down the hierarchy and closer to the tactical edge, the communication network exhibits a stronger response to external events. The tactical edge is commonly understood to be the part of the organization that interacts directly with the operating environment (Alberts & Hayes, 2003), and so our findings align with the intuitive understanding of Army organization. The correspondence between position in the hierarchy and sensitivity to external events can be further tested by exploring similar scenario-based data in even lower echelons (battalion, company, platoon...) to see how their communications are affected by exercise events.

Our results have considerable potential for application to military operations, for example in the sphere of bandwidth usage and allocation. Relevant battlefield events are already currently recorded and tracked through SIGACTs (significant activity reports). Understanding the different relationships at different echelons between these events and the resulting communications within the Mission Command network would make it possible to use the SIGACT record to make real-time predictions of communications volume. This could in turn allow for more targeted dynamic bandwidth allocation and a more efficient and agile Mission Command communications network.

Acknowledgement

Research was sponsored by the Army Research Laboratory and was accomplished under Cooperative Agreement Number W911NF-12-2-0019. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Laboratory or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

References

- Alberts, D. S., & Hayes, R. E. (2003). *Power to the edge: Command... control... in the information age*. Office of the Assistant Secretary of Defense Washington DC Command and Control Research Program (CCRP).
- Ancona, D. G., & Caldwell, D. F. (1992). Bridging the boundary: External activity and performance in organizational teams. *Administrative science quarterly*, 37, 634-665.
- Assmus, G., Farley, J. U., & Lehmann, D. R. (1984). How advertising affects sales: Meta-analysis of econometric results. *Journal of Marketing Research*, 21, 65-74.
- Barabási, A. L. (2010). *Bursts: The hidden pattern behind everything we do*. EP Dutton.
- Barsalou, L. W. (2010). Grounded cognition: Past, present, and future. *Topics in Cognitive Science*, 2, 716-724.
- Carley, K. M. (2003, November). Dynamic network analysis. In *Dynamic social network modeling and analysis: Workshop summary and papers* (pp. 133-145). Committee on Human Factors, National Research Council.
- Castells, M. (1996). The Rise of the Network Society, Volume I of The Information Age: Economy, Society and Culture. *Massachusetts: Blackwell Publishing*.
- Clark (1997) *Being There: Putting Brain, Body and World Together Again*. MIT Press: Cambridge, Massachusetts.
- Contractor, N. S., Wasserman, S., & Faust, K. (2006). Testing multitheoretical, multilevel hypotheses about organizational networks: An analytic framework and empirical example. *Academy of Management Review*, 31, 681-703.
- Danowski, J. A., & Edison-Swift, P. (1985). Crisis effects on intraorganizational computer-based communication. *Communication Research*, 12, 251-270.
- Diesner, J., Frantz, T. L., & Carley, K. M. (2005). Communication networks from the Enron email corpus "It's always about the people. Enron is no different". *Computational & Mathematical Organization Theory*, 11(3), 201-228.
- Hackman, J.R. (2008). What Is This Thing Called Leadership? In *Harvard Business School Centennial Colloquium on "Leadership: Advancing an Intellectual Discipline," June* (pp. 10-11).
- Hackman, J.R. & Katz, N. (2010). Group behavior and performance. In S.T. Fiske, D.T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5th ed.), New York: Wiley, 1208-1251.

- Kilduff, M., & Brass, D. J. (2010). Organizational social network research: Core ideas and key debates. *The Academy of Management Annals*, 4, 317-357.
- Leone, R. P. (1983). Modeling sales-advertising relationships: An integrated time series-econometric approach. *Journal of Marketing Research*, 20, 291-295.
- McGrath, J. E., & Kravitz, D. A. (1982). Group research. *Annual Review of Psychology*, 33, 195-230.
- Monge, P. R., & Contractor, N. S. (2001). Emergence of communication networks. *The new handbook of organizational communication: Advances in theory, research, and methods*, 440-502.
- National Research Council. *Network Science*. Washington, DC: The National Academies Press, 2005.
- Parsons, L. J., & Schultz, R. L. (1976). *Marketing models and econometric research*. New York, Amsterdam: North-Holland Publishing Company.
- Potterfield, Thomas. "The Business of Employee Empowerment: Democracy and Ideology in the Workplace." Quorum Books, 1999, p. 6
- Roby, T. B., & Lanzetta, J. T. (1958). Considerations in the analysis of group tasks. *Psychological Bulletin*, 55, 88.
- Roth, J.T., Hrits, R.J., & McGill, D.W. (1984). Model of team organization and behavior and team description method (Research Note 84-129). U.S. Army Research Institute for the Behavioral and Social Sciences.
- Shah, P. P. (2000). Network destruction: The structural implications of downsizing. *Academy of Management Journal*, 43, 101-112.
- Simon, H. A. (1996). *The sciences of the artificial*. MIT press.
- Snijders, Tom A.B. (2005). Models for Longitudinal Network Data. Chapter 11 (pp. 215{247) in P. Carrington, J. Scott, and S. Wasserman (Eds.), *Models and methods in social network analysis*. New York: Cambridge University Press.
- Vázquez, A., Oliveira, J. G., Dezsö, Z., Goh, K. I., Kondor, I., & Barabási, A. L. (2006). Modeling bursts and heavy tails in human dynamics. *Physical Review E*, 73, 036127.



U.S. Army Research, Development and Engineering Command

Time Series Modeling of Army Mission Command Communication Networks: An Event-Driven Analysis

ARL

TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Laura Marusich
Norbou Buchler
U.S. Army Research Laboratory
ICCRTS 2013

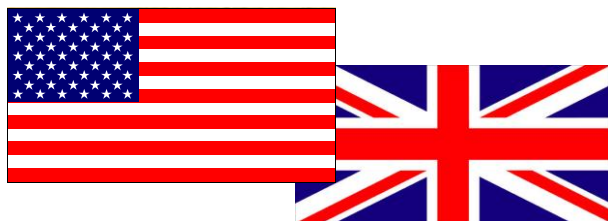
Motivation

- **Transition to Network-Enabled Operations**
- **Vast number of potential collaborations**
- **Vast amount of available information**
- **C2 networks increasingly dependent upon robust human interactions for timely decision-making**
- **Can we predict patterns of communication occurring across time?**

- **Examine the communication record of a fully-networked Army coalition command and control organization**
 - Two echelons of command at the Division and Brigade levels over a two-week military scenario exercise
- **Use communication record to construct social network at each time point of the exercise**
- **Use time series analysis methods to predict number of communication links at a given time point**

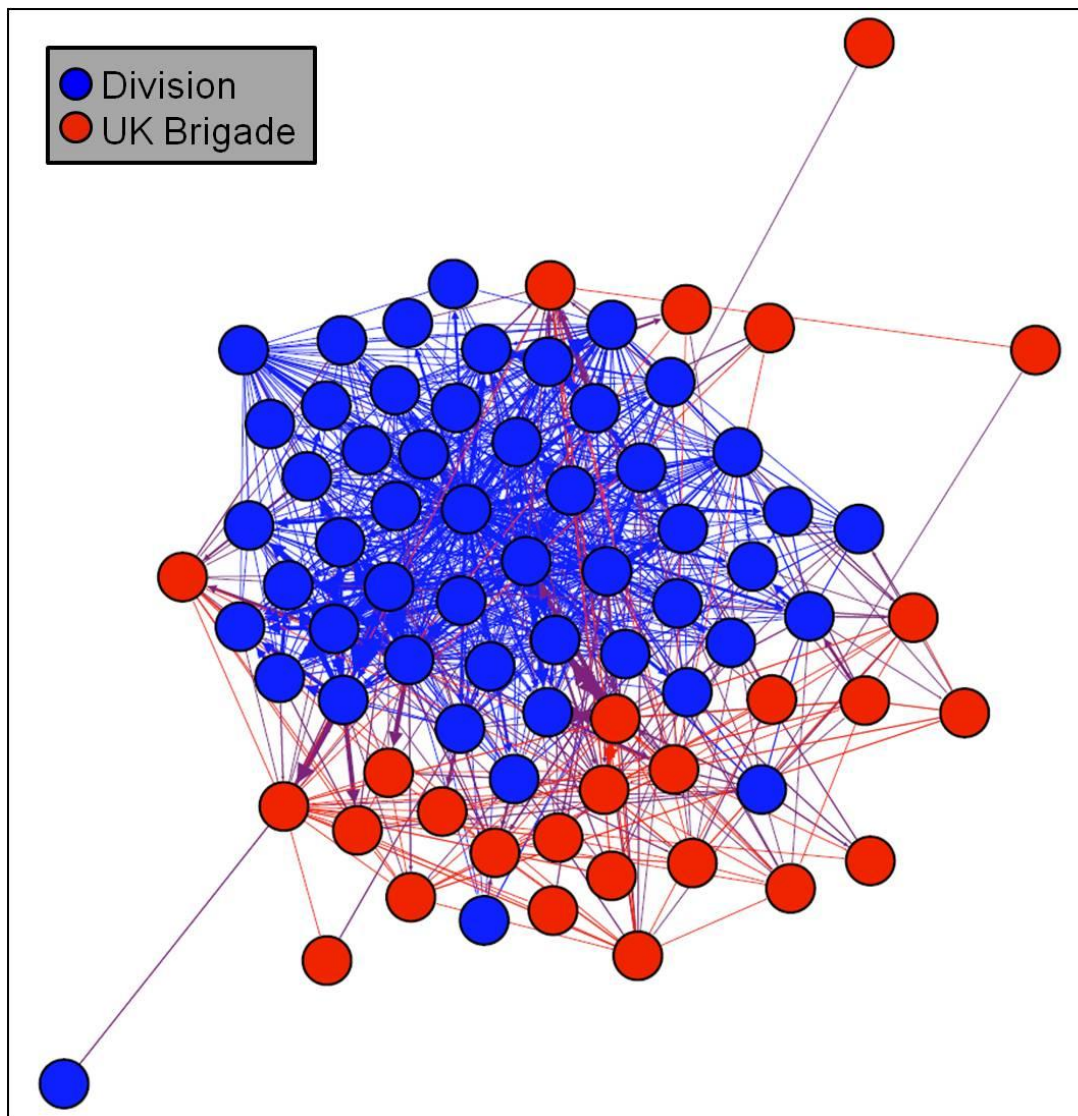
U.S. Army simulation-based training event

- Two week experiment/exercise at Mission Command Battle Laboratory, Ft. Leavenworth
 - Week 1 – lethal operations
 - Week 2 – civil-military operations
- Participating Groups
 - U.S. Division HQ
 - U.K. Brigade Combat Team
 - Two U.S. Brigade Combat Teams (Not fully staffed)
- Focus: U.S./U.K interoperability

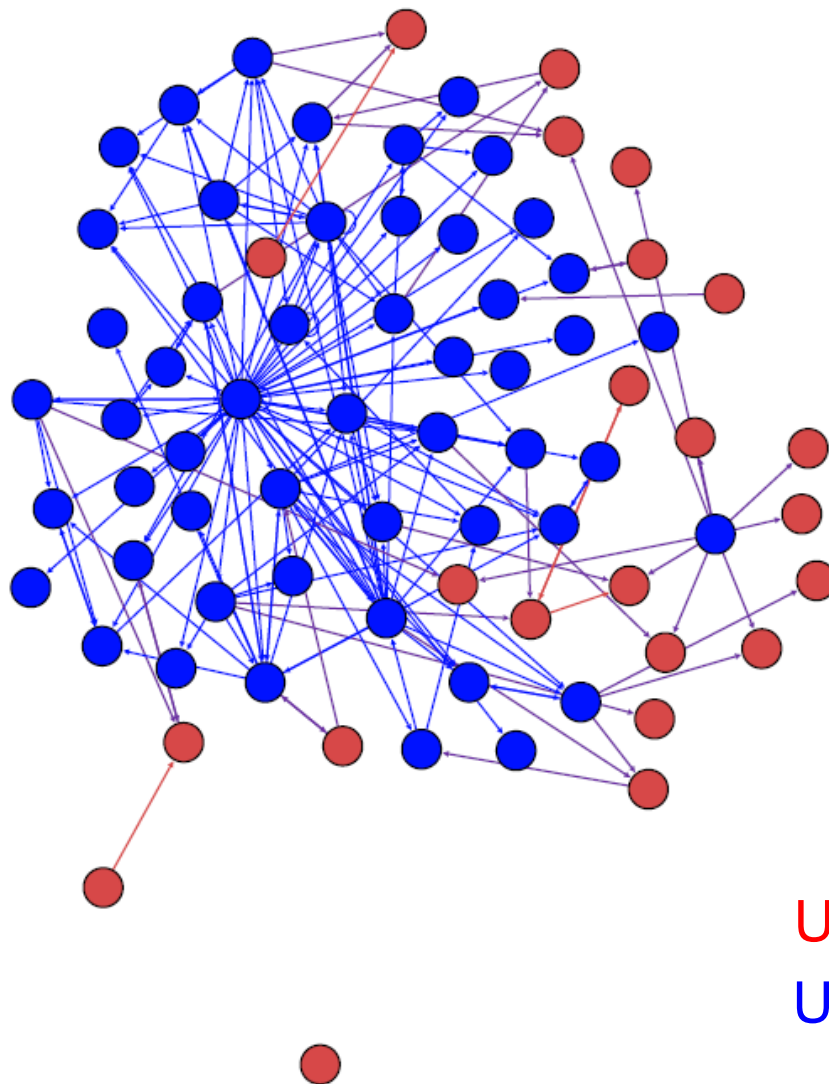


- Participants primarily used telephone (VoIP) and email to communicate
- Collected three pieces of information:
 - Sender
 - Receiver
 - Timestamp
- Aggregated the communications into intervals of one hour
- Constructed communications networks for each interval

Overall Network



Network by Hour

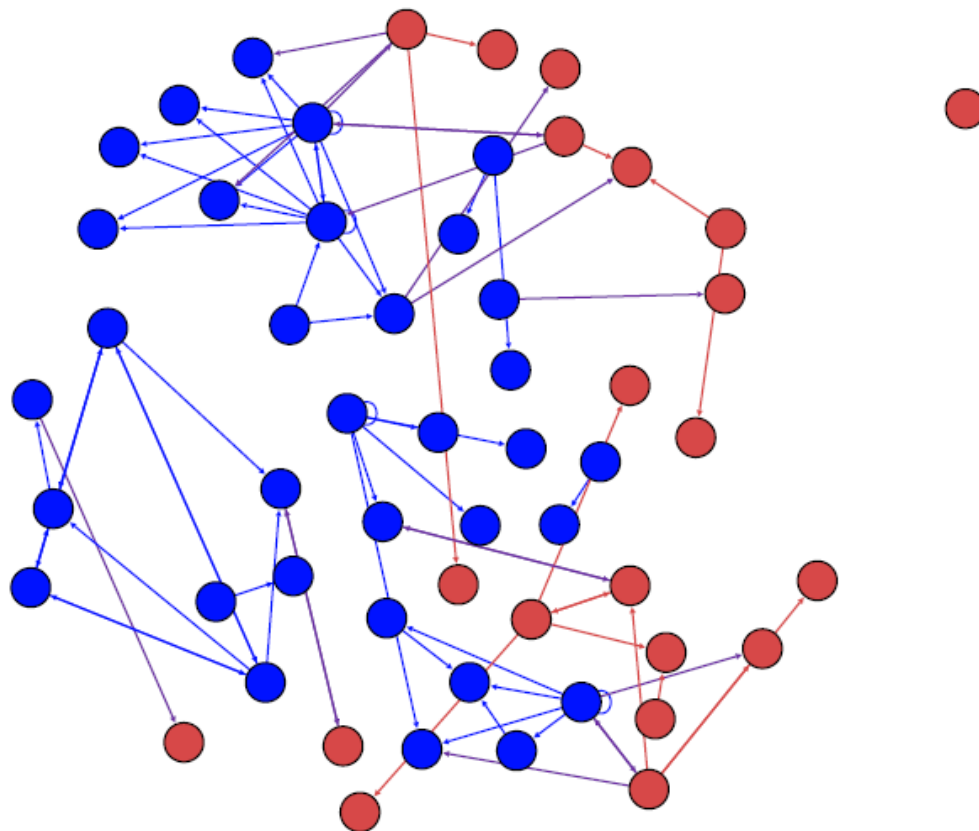


Hour 1

U.K. Brigade

U.S. Division

Network by Hour

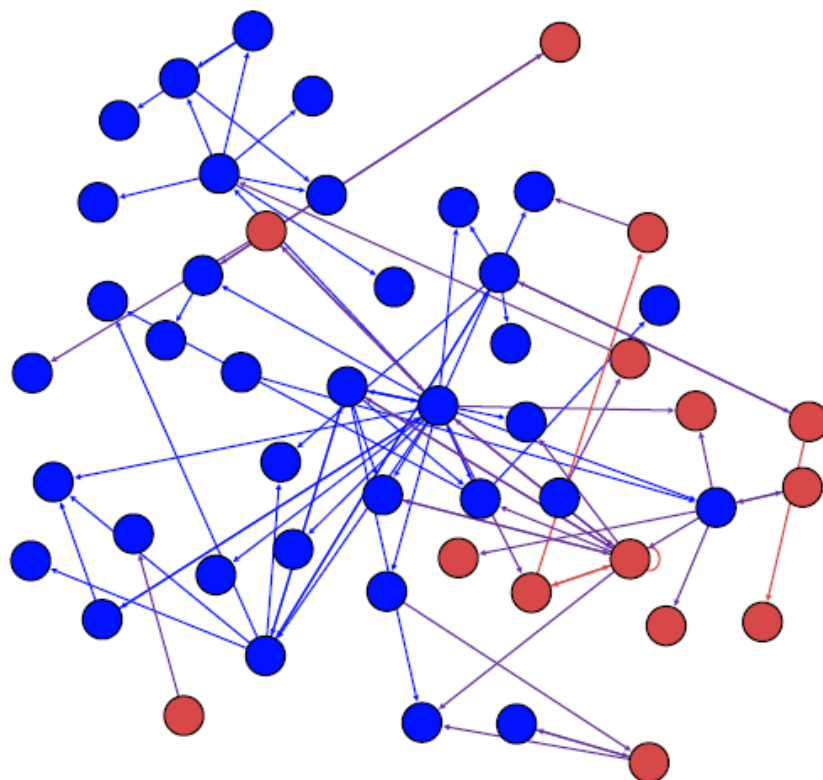


Hour 2

U.K. Brigade

U.S. Division

Network by Hour

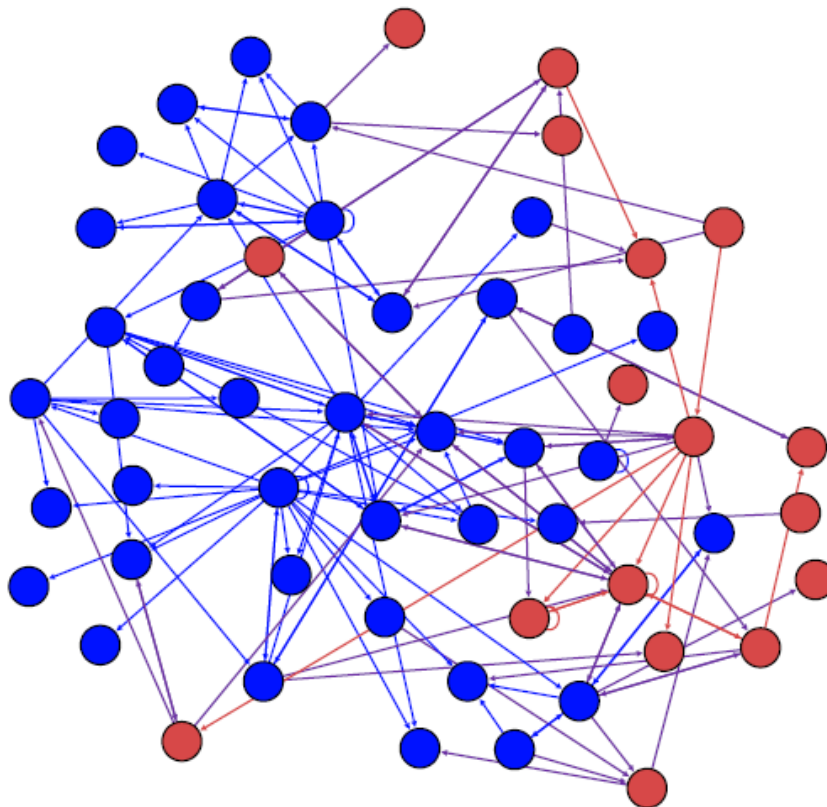


Hour 3

U.K. Brigade

U.S. Division

Network by Hour

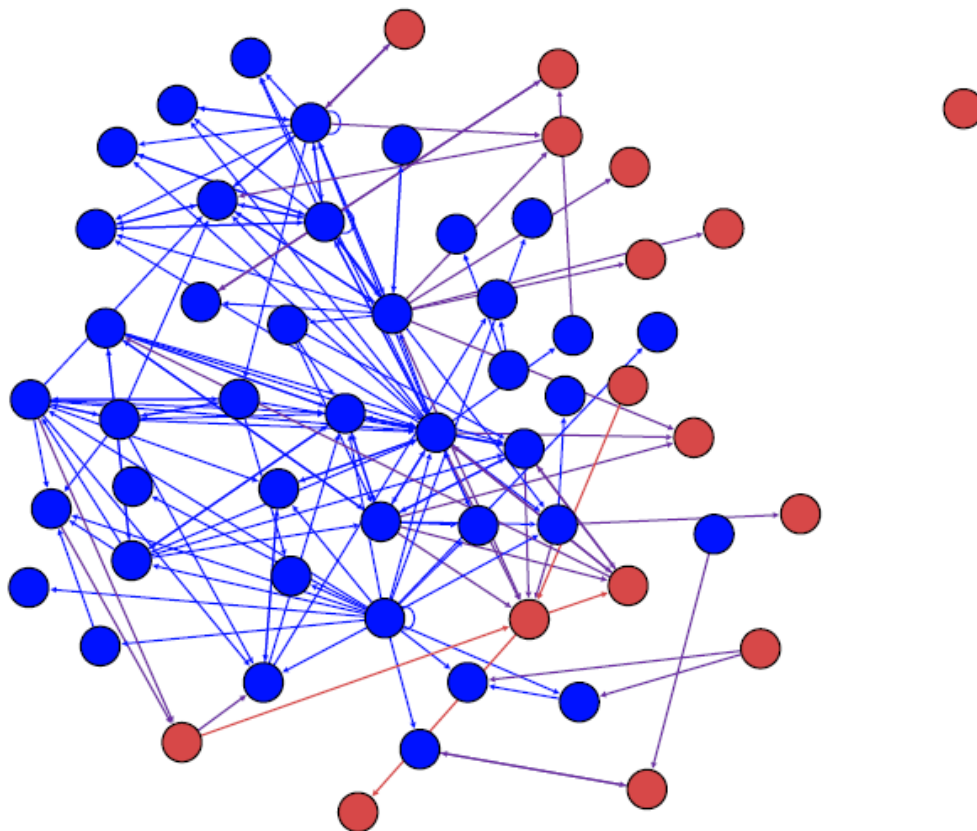


Hour 4

U.K. Brigade

U.S. Division

Network by Hour

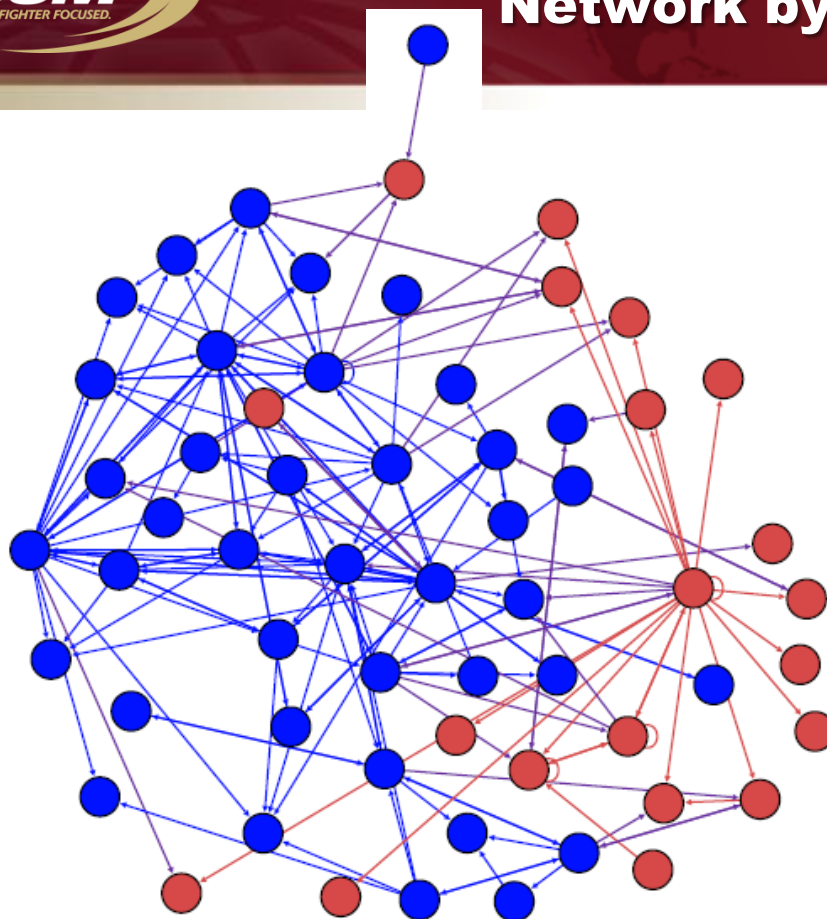


Hour 5

U.K. Brigade

U.S. Division

Network by Hour

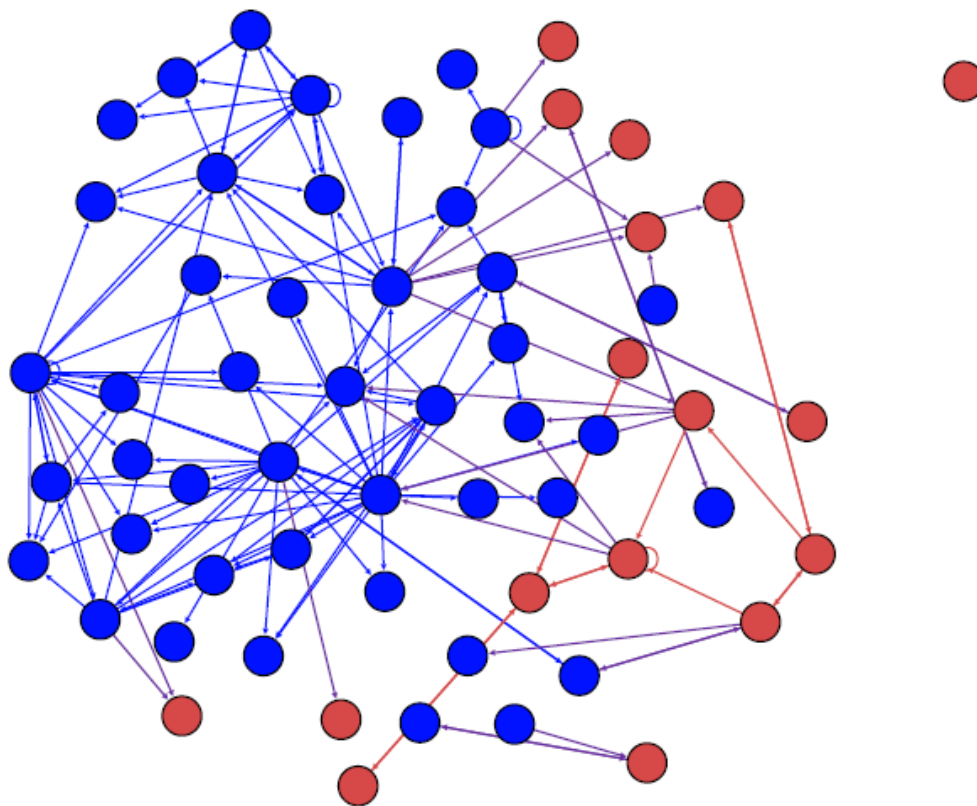


Hour 6

U.K. Brigade

U.S. Division

Network by Hour

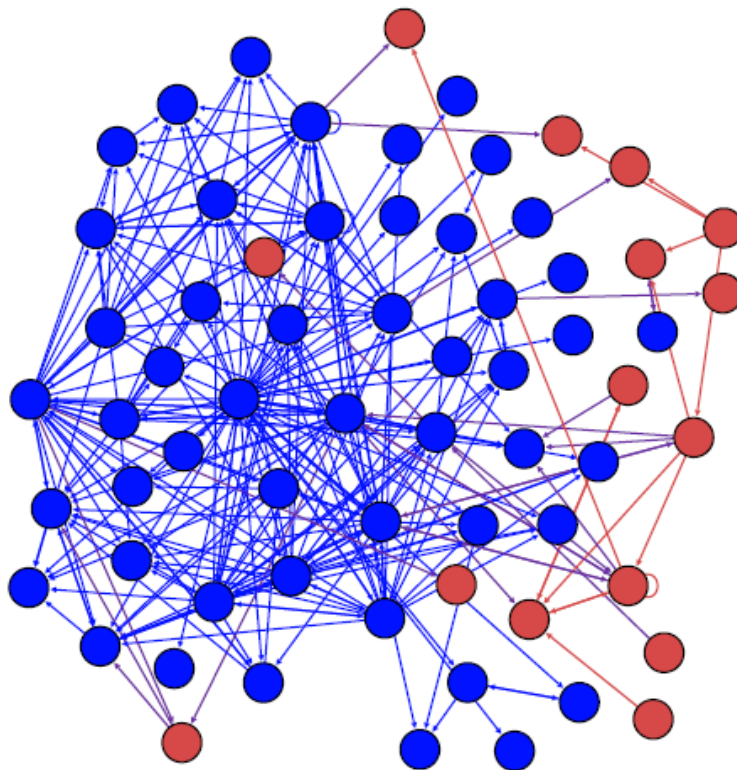


Hour 7

U.K. Brigade

U.S. Division

Network by Hour

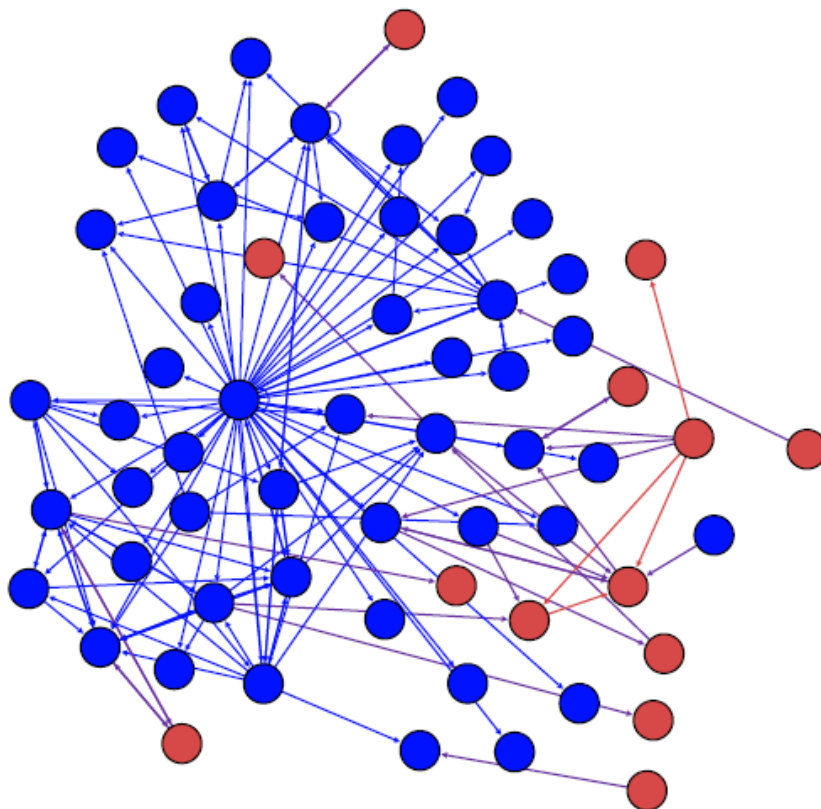


Hour 8

U.K. Brigade

U.S. Division

Network by Hour

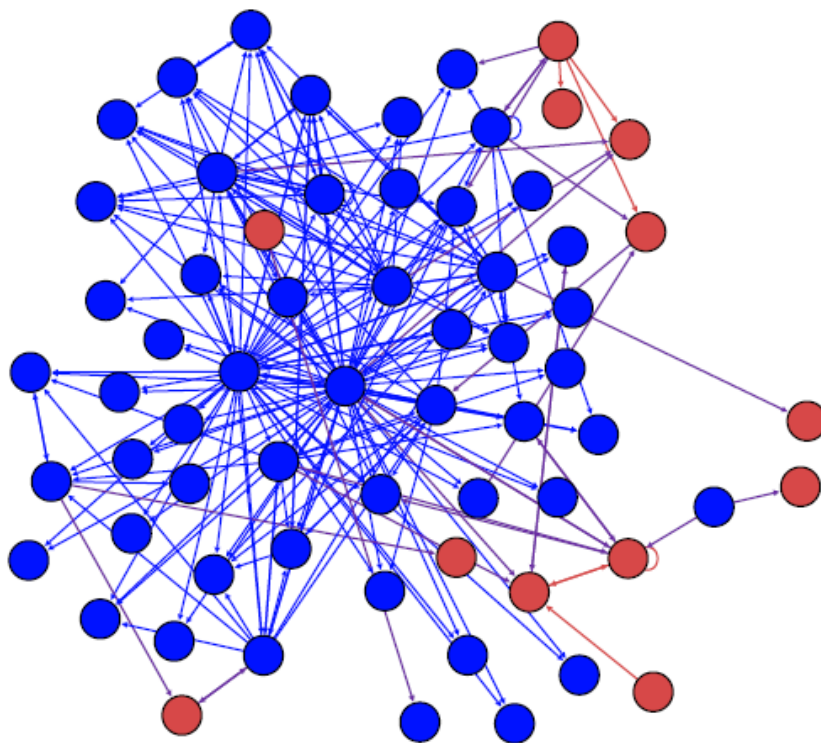


Hour 9

U.K. Brigade

U.S. Division

Network by Hour

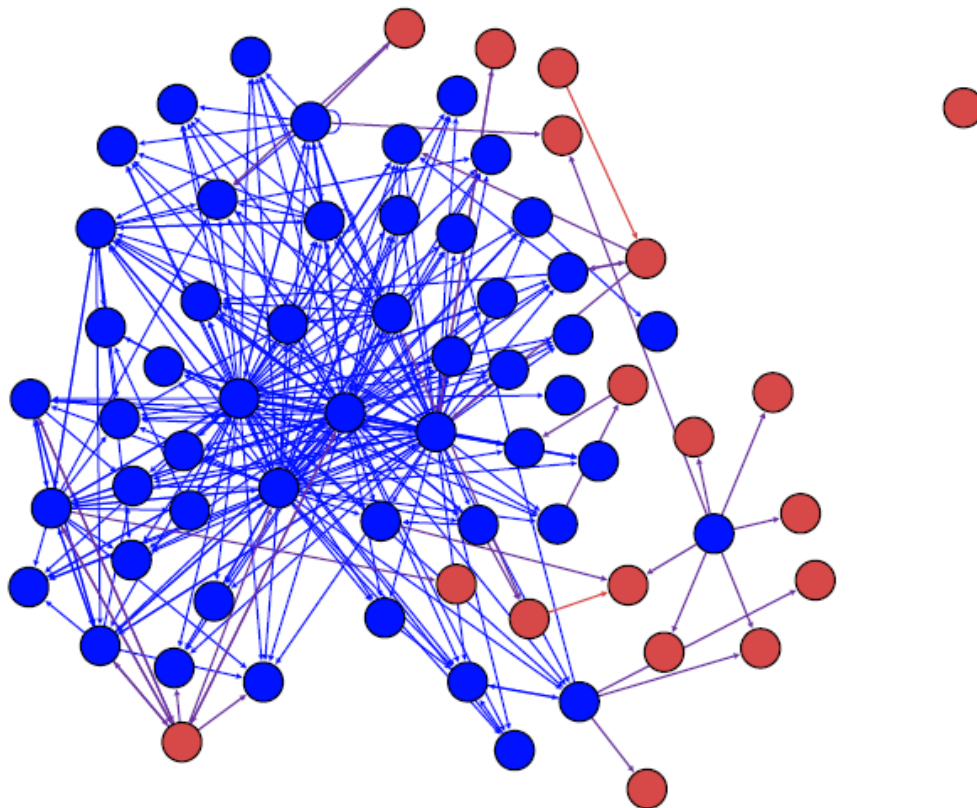


Hour 10

U.K. Brigade

U.S. Division

Network by Hour

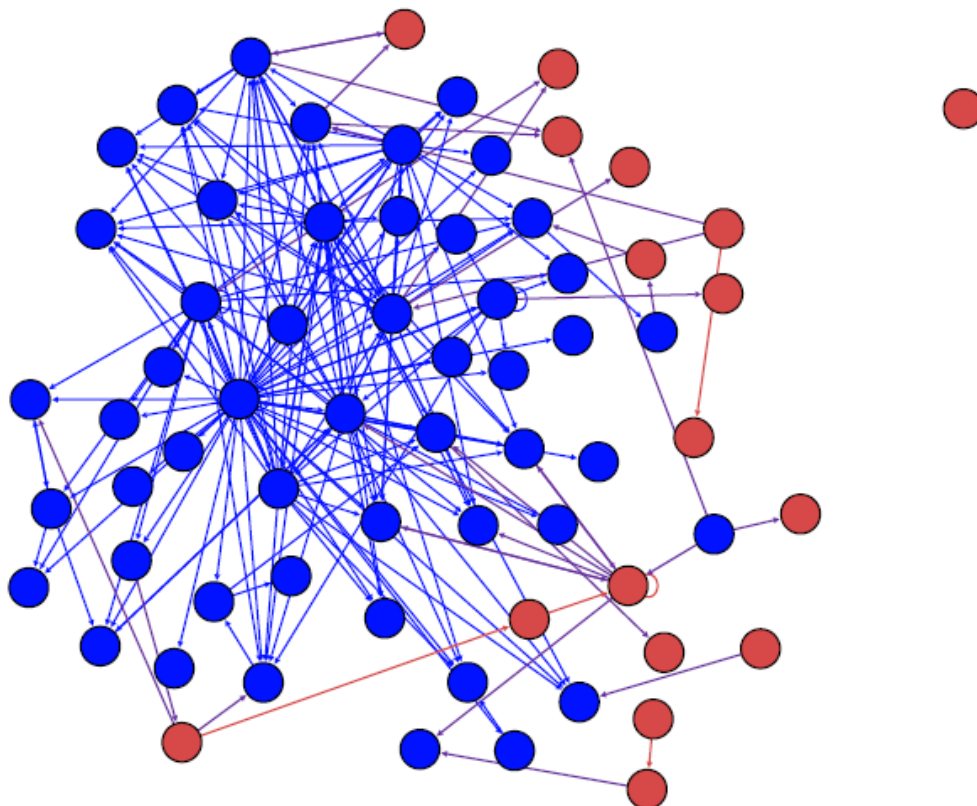


Hour 11

U.K. Brigade

U.S. Division

Network by Hour

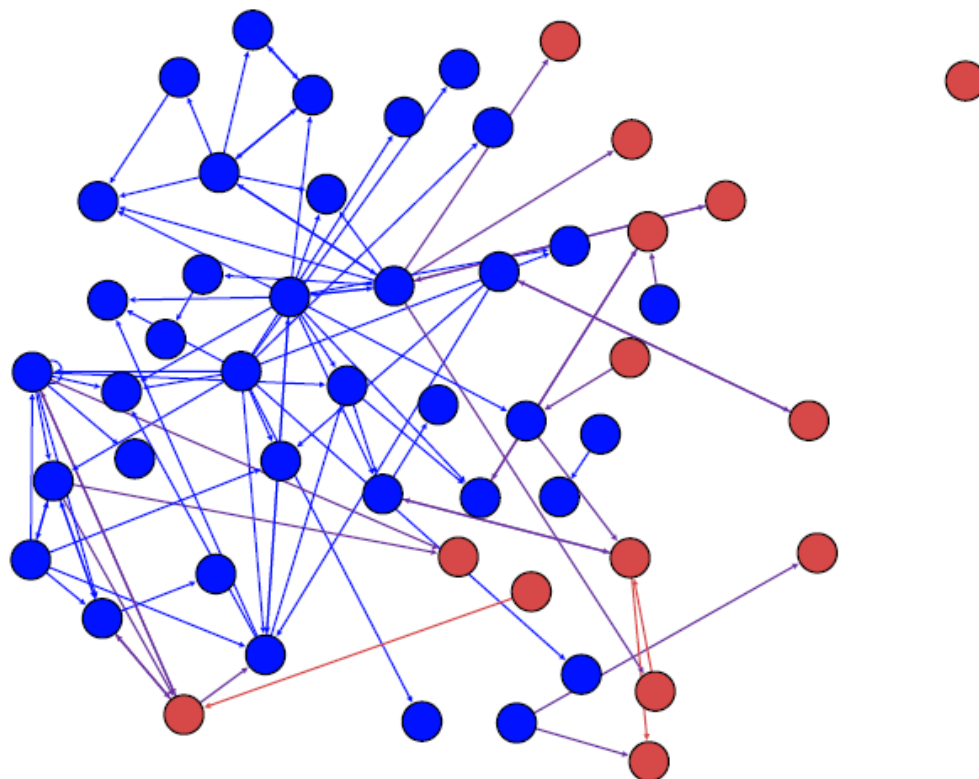


Hour 12

U.K. Brigade

U.S. Division

Network by Hour

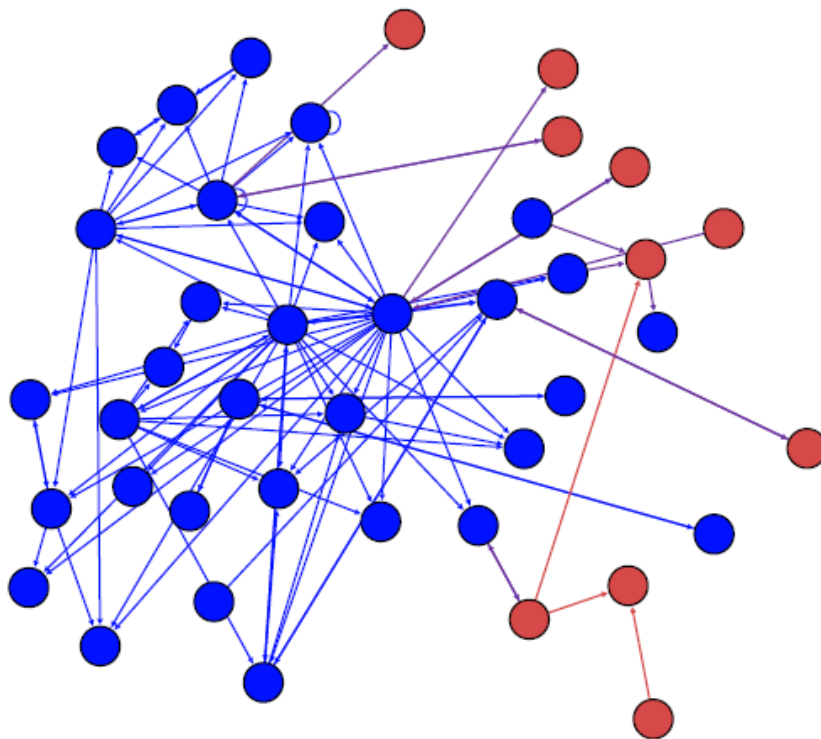


Hour 13

U.K. Brigade

U.S. Division

Network by Hour

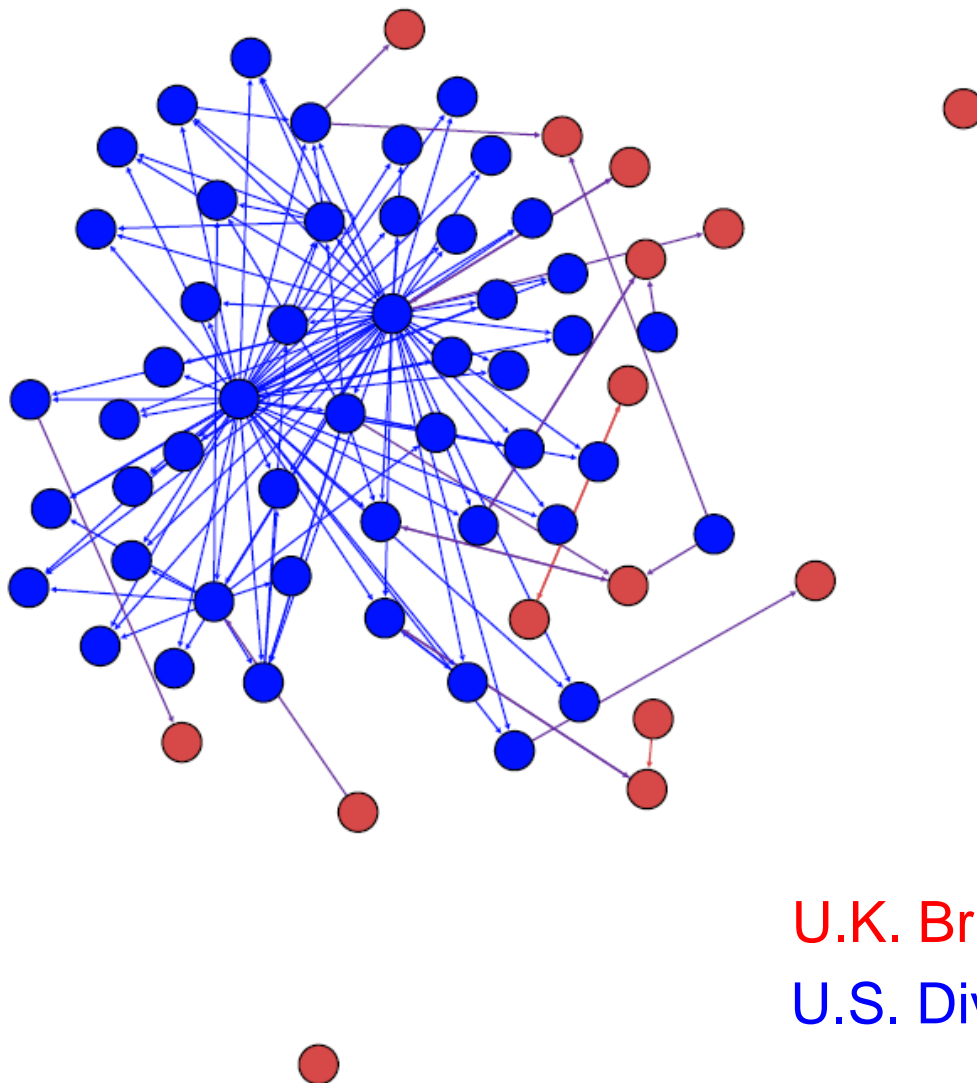


Hour 14

U.K. Brigade

U.S. Division

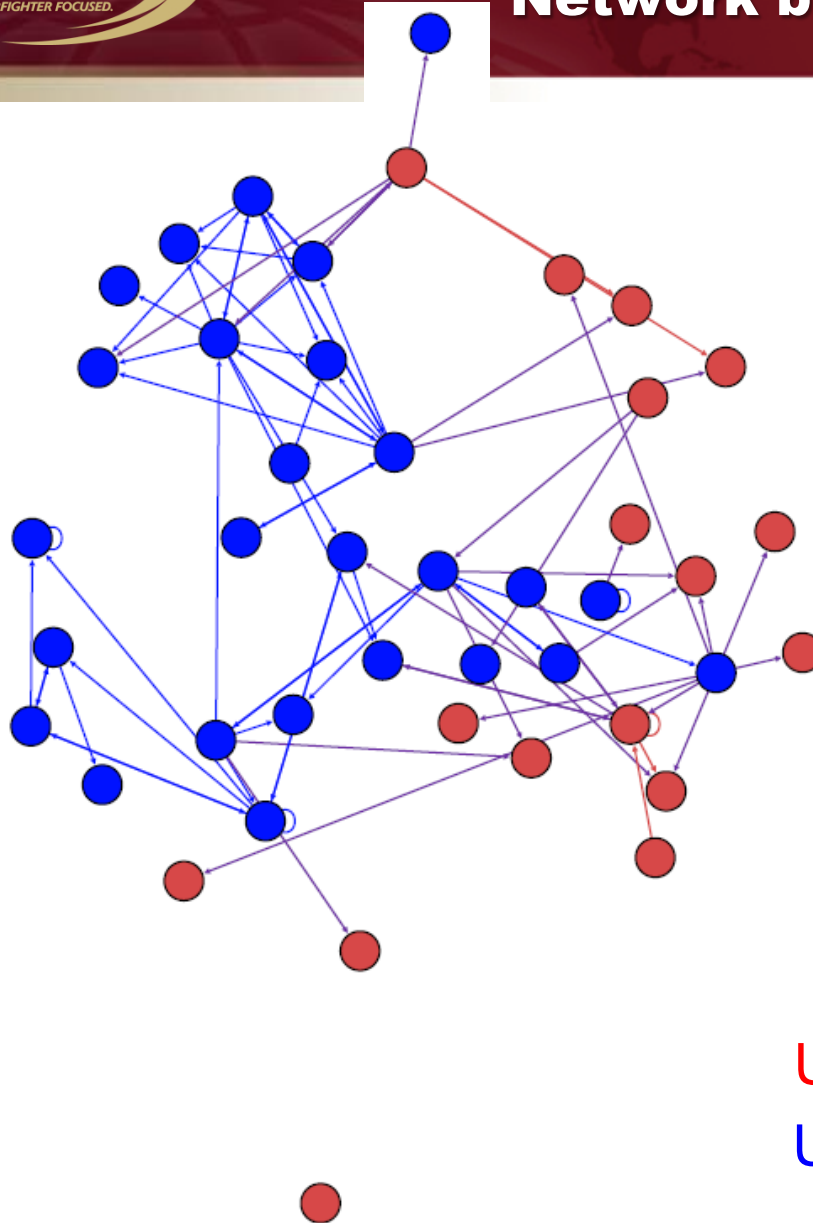
Network by Hour



Hour 15

U.K. Brigade

U.S. Division



Hour 16

U.K. Brigade

U.S. Division

● Internal Links (1)

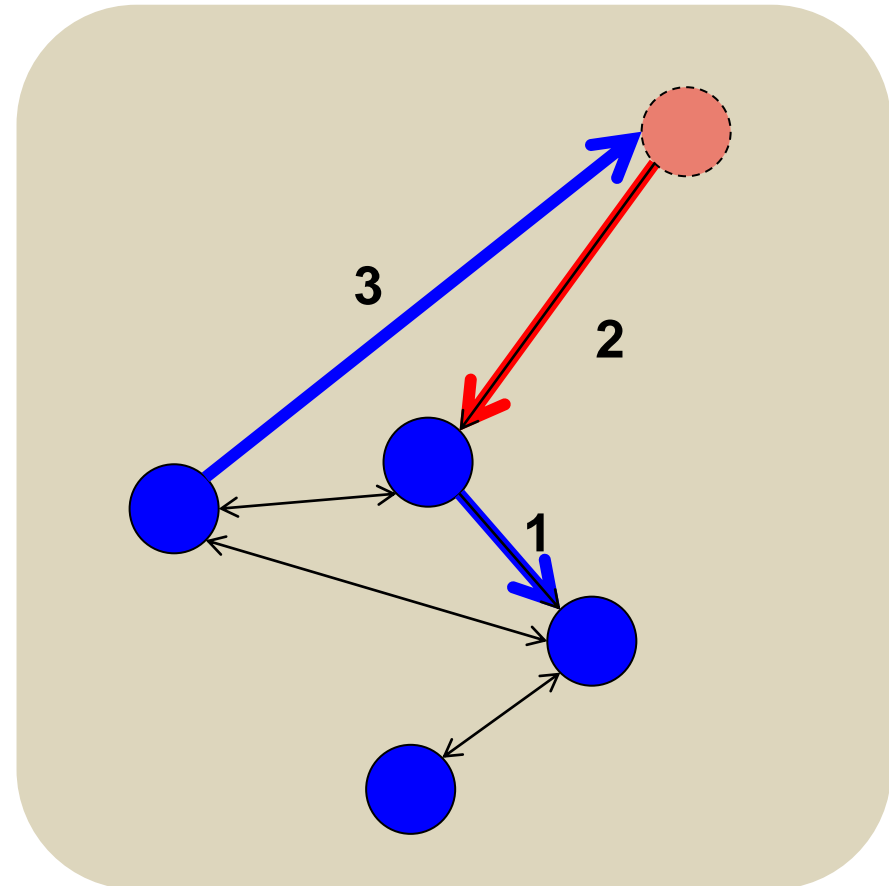
- e.g. Link from a division member to another division member

● External Links IN (2)

- e.g. Link from member of a different group to a division member

● External Links OUT (3)

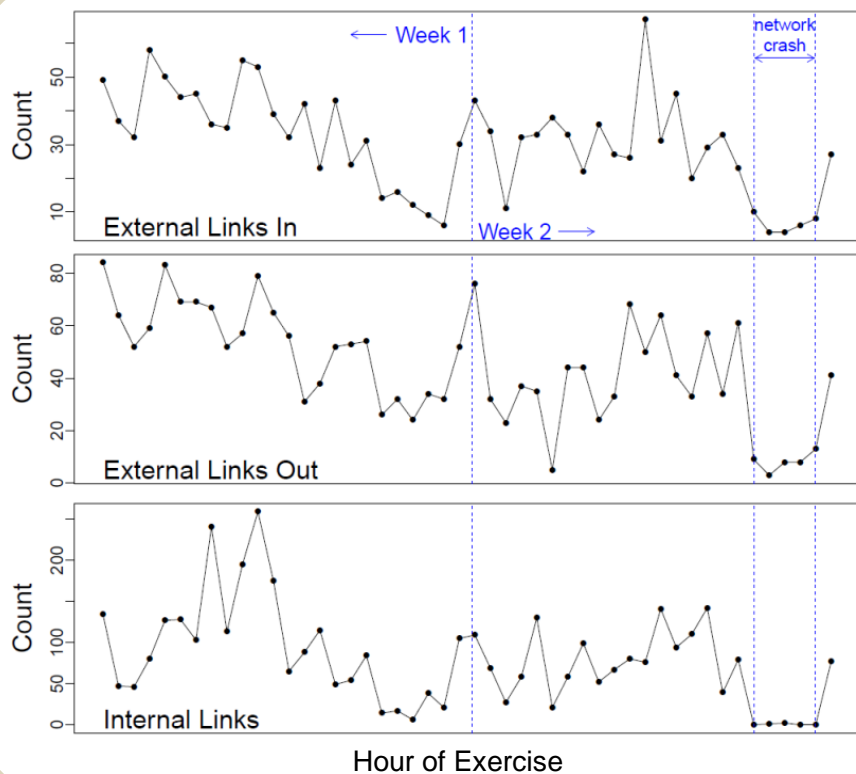
- e.g. Link from division member to member of a different group



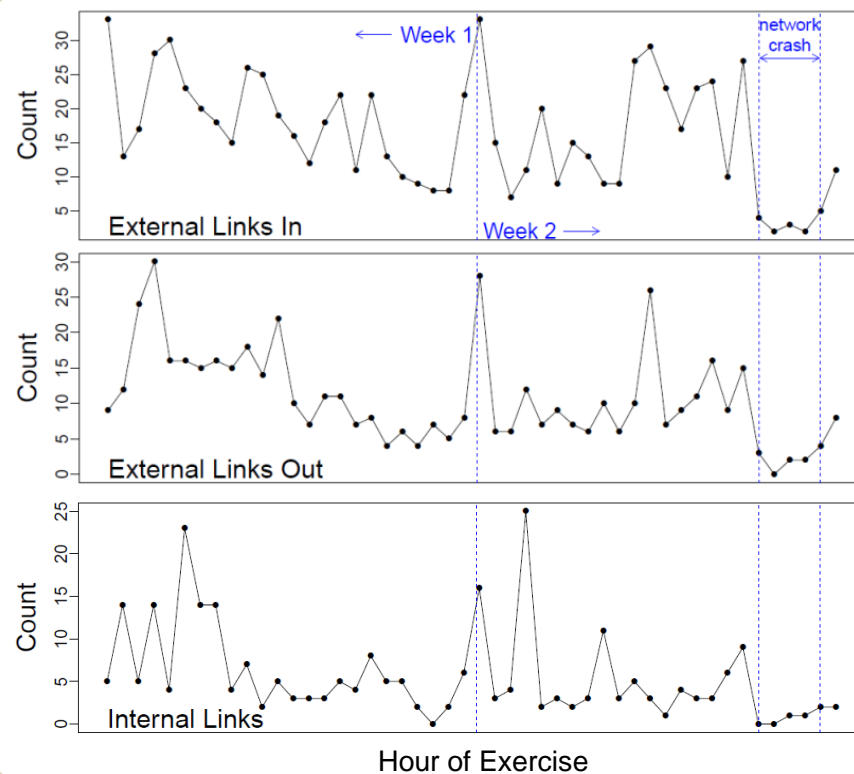
- **Training scenario generates overlapping series of event-driven tasks, requiring coordination among the C2 staff.**
- **Examples:**
 - Improvised explosive device (IED) in the area of operations
 - Civil demonstration
 - Intelligence report of enemy activity
- **Participants adhere to formalized military work routines and processes**
- **Events are time-sensitive and may have adverse cascading effects if not addressed in a timely manner.**
- **Participants construct specific, detailed material products (e.g. regular reports to the Commander from the staff)**

Time Series

Division

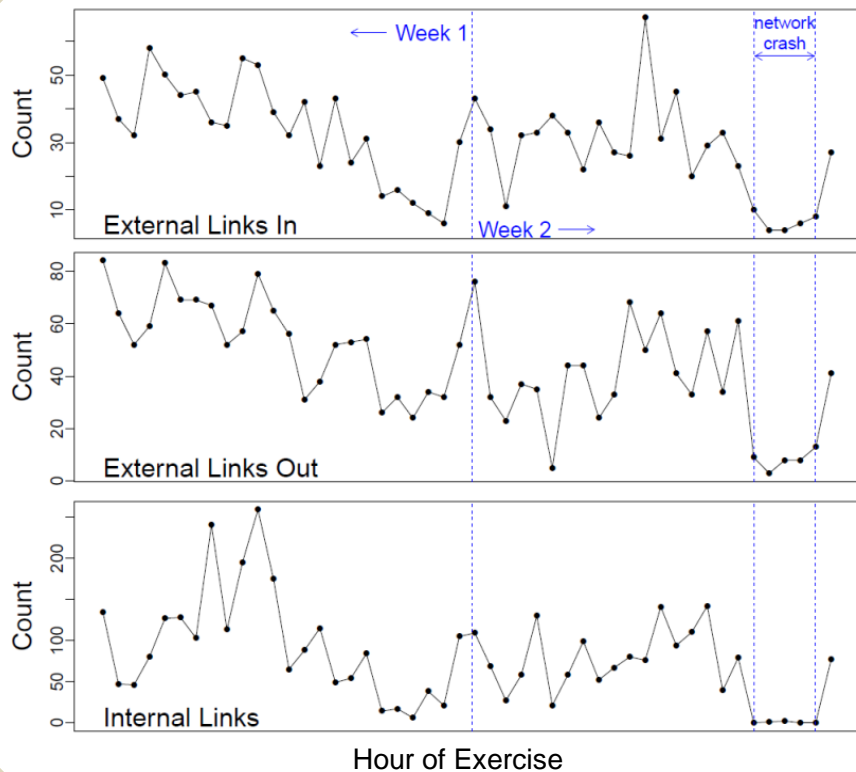


UK Brigade

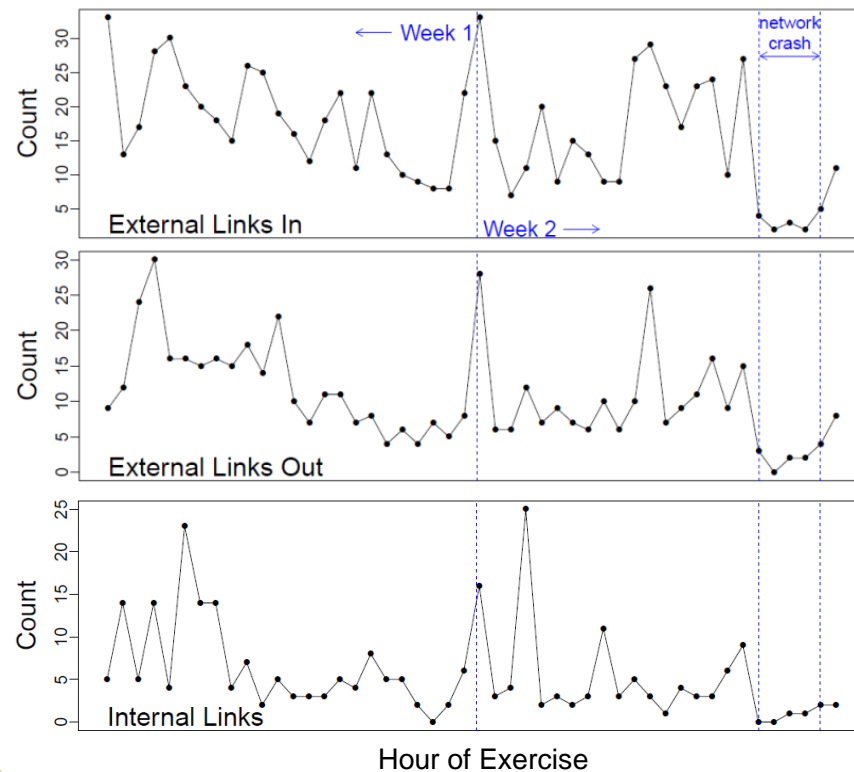


Time Series

Division

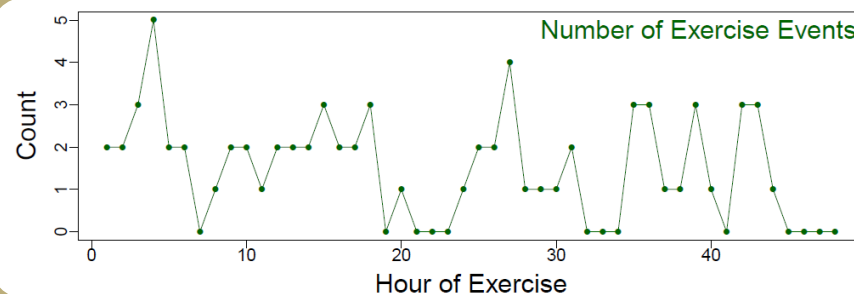


UK Brigade



Communications

Events



Division

Autoregressive Moving Average (ARMA) Model

- AR (1)

$$y_t = \alpha_0 + \alpha_1 y_{t-1}$$

- Predicts only on the basis of internal dynamics
- Only the immediately previous value of y is used to predict the current value

UK Brigade

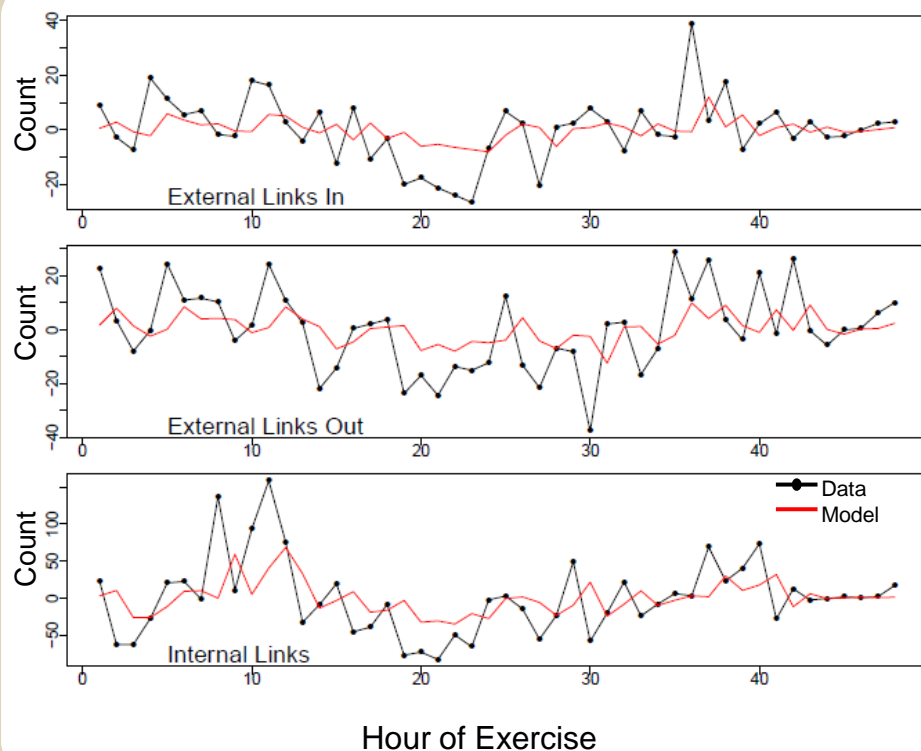
Lag Dependent Variable (LDV) Model

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \beta x_t$$

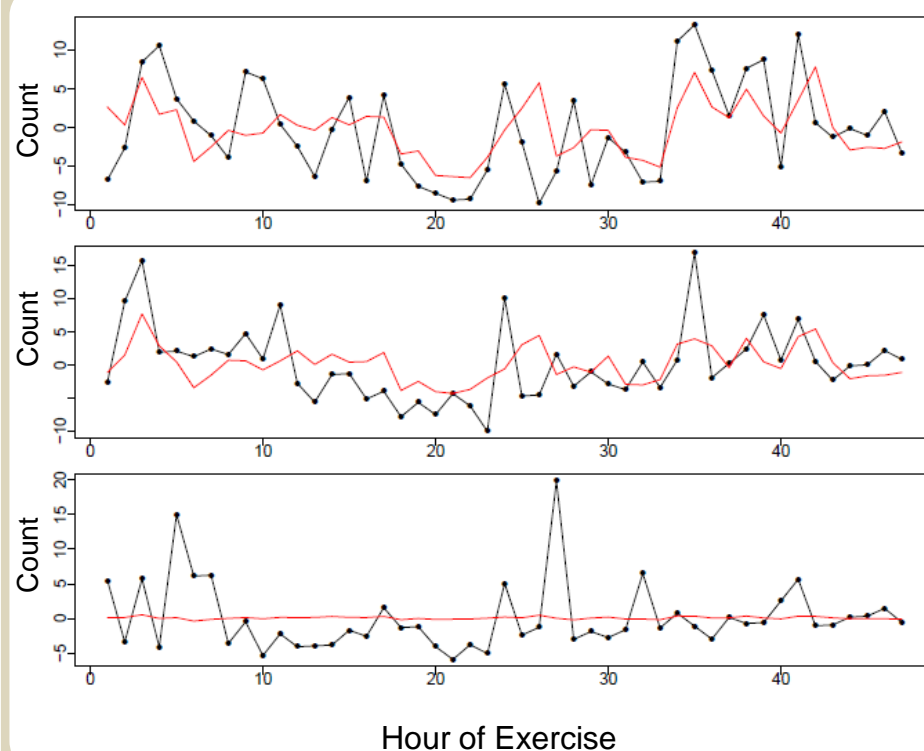
- Predictors:
 - Current number of events - x
 - Previous value of y
- Predicts both long- and short-run effects of changes in the number of events

Model Predictions

Division ARMA Model



UK Brigade LDV Model



- **Brigade communications best predicted by:**
 - Scenario events
 - Past communications
- **Division communications best predicted by:**
 - Past communications alone

- **Results suggest a correspondence between position in the hierarchy and sensitivity to external events**
 - Communication networks closer to the tactical edge exhibit a stronger response to external events.
- **Future work:**
 - Explore similar scenario-based data in even lower echelons (battalion, company, platoon...) to see if this relationship holds.
- **Application: Bandwidth usage and allocation.**
 - Relevant battlefield events are already currently recorded and tracked through significant activity reports (SIGACTs).
 - This SIGACT record could be used to make real-time predictions of communications volume.
 - Such predictions could in turn allow for more targeted dynamic bandwidth allocation.